

BC #5
PCT09

P.5

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,352

DATE: 10/30/2001

TIME: 14:31:45

Input Set : A:\gje-70.txt

Output Set: N:\CRF3\10302001\I868352.raw

ENTERED

```

5 <110> APPLICANT: Hughes, Martin J G
7   Santangelo, Joseph D
9   Lane, Jonathan D
11  Feldman, Robert
13  Moore, Joanne C
15  Dobson, Richard J
17  Everest, Paul
19  Dougan, Gordon
21  Wilson, Rebecca K
25 <120> TITLE OF INVENTION: Genes and Proteins, and Their Use
29 <130> FILE REFERENCE: GJE-70
33 <140> CURRENT APPLICATION NUMBER: US 09/868,352
35 <141> CURRENT FILING DATE: 2001-06-15
39 <160> NUMBER OF SEQ ID NOS: 35
43 <170> SOFTWARE: PatentIn version 3.1
47 <210> SEQ ID NO: 1
49 <211> LENGTH: 587
51 <212> TYPE: DNA
53 <213> ORGANISM: Streptococcus agalactiae
57 <220> FEATURE:
59 <221> NAME/KEY: CDS
61 <222> LOCATION: (1)..(582)
63 <223> OTHER INFORMATION:
67 <400> SEQUENCE: 1
68 atg atc cca gta gta atc gaa caa aca agt cgt ggt gaa cgt tct tat      48
69 Met Ile Pro Val Val Ile Glu Gln Thr Ser Arg Gly Glu Arg Ser Tyr
70 1           5           10           15
72 gat att tac tca cgt ctt tta aaa gat cgt att att atg ttg aca ggc      96
73 Asp Ile Tyr Ser Arg Leu Leu Lys Asp Arg Ile Ile Met Leu Thr Gly
74           20           25           30
76 caa gtt gag gat aat atg gcc aat agt atc att gca cag tta ttg ttt      144
77 Gln Val Glu Asp Asn Met Ala Asn Ser Ile Ile Ala Gln Leu Leu Phe
78           35           40           45
80 ctc gat gca caa gat aat aca aag gat att tac ctt tat gtc aat aca      192
81 Leu Asp Ala Gln Asp Asn Thr Lys Asp Ile Tyr Leu Tyr Val Asn Thr
82           50           55           60
84 cca ggt ggt tca gta tcg gct gga ctt gct att gtg gac acc atg aac      240
85 Pro Gly Gly Ser Val Ser Ala Gly Leu Ala Ile Val Asp Thr Met Asn
86 65           70           75           80
88 ttc att aaa tcg gac gta cag acg att gtt atg ggg atg gct gct tcg      288
89 Phe Ile Lys Ser Asp Val Gln Thr Ile Val Met Gly Met Ala Ala Ser
90           85           90           95
92 atg gga acc att att gct tca agt ggt gct aaa gga aaa cgt ttt atg      336
93 Met Gly Thr Ile Ile Ala Ser Ser Gly Ala Lys Gly Lys Arg Phe Met
94           100          105          110
96 tta ccg aat gca gaa tat atg atc cac caa cca atg ggc gga aca ggc      384
97 Leu Pro Asn Ala Glu Tyr Met Ile His Gln Pro Met Gly Gly Thr Gly

```

RAW SEQUENCE LISTING

DATE: 10/30/2001

PATENT APPLICATION: US/09/868,352

TIME: 14:31:45

Input Set : A:\gje-70.txt

Output Set: N:\CRF3\10302001\I868352.raw

```

98          115          120          125
100 gga ggt aca cag caa tct gat atg gct atc gct gct gag cat ctt tta      432
101 Gly Gly Thr Gln Gln Ser Asp Met Ala Ile Ala Ala Glu His Leu Leu
102          130          135          140
104 aaa acg cgt cat act tta gaa aaa atc tta gct gat aat tct ggt caa      480
105 Lys Thr Arg His Thr Leu Glu Lys Ile Leu Ala Asp Asn Ser Gly Gln
106 145          150          155          160
108 tct att gaa aaa gtc cat gat gat gca gag cgt gat cgt tgg atg agt      528
109 Ser Ile Glu Lys Val His Asp Asp Ala Glu Arg Asp Arg Trp Met Ser
110          165          170          175
112 gct caa gaa aca ctt gat tat ggc ttt att gat gaa atc atg gct aat      576
113 Ala Gln Glu Thr Leu Asp Tyr Gly Phe Ile Asp Glu Ile Met Ala Asn
114          180          185          190
116 aat gaa taagg      587
117 Asn Glu
121 <210> SEQ ID NO: 2
123 <211> LENGTH: 194
125 <212> TYPE: PRT
127 <213> ORGANISM: Streptococcus agalactiae
131 <400> SEQUENCE: 2
133 Met Ile Pro Val Val Ile Glu Gln Thr Ser Arg Gly Glu Arg Ser Tyr
134 1          5          10          15
137 Asp Ile Tyr Ser Arg Leu Leu Lys Asp Arg Ile Ile Met Leu Thr Gly
138          20          25          30
141 Gln Val Glu Asp Asn Met Ala Asn Ser Ile Ile Ala Gln Leu Leu Phe
142          35          40          45
145 Leu Asp Ala Gln Asp Asn Thr Lys Asp Ile Tyr Leu Tyr Val Asn Thr
146          50          55          60
149 Pro Gly Gly Ser Val Ser Ala Gly Leu Ala Ile Val Asp Thr Met Asn
150 65          70          75          80
153 Phe Ile Lys Ser Asp Val Gln Thr Ile Val Met Gly Met Ala Ala Ser
154          85          90          95
157 Met Gly Thr Ile Ile Ala Ser Ser Gly Ala Lys Gly Lys Arg Phe Met
158          100          105          110
161 Leu Pro Asn Ala Glu Tyr Met Ile His Gln Pro Met Gly Gly Thr Gly
162          115          120          125
165 Gly Gly Thr Gln Gln Ser Asp Met Ala Ile Ala Ala Glu His Leu Leu
166          130          135          140
169 Lys Thr Arg His Thr Leu Glu Lys Ile Leu Ala Asp Asn Ser Gly Gln
170 145          150          155          160
173 Ser Ile Glu Lys Val His Asp Asp Ala Glu Arg Asp Arg Trp Met Ser
174          165          170          175
177 Ala Gln Glu Thr Leu Asp Tyr Gly Phe Ile Asp Glu Ile Met Ala Asn
178          180          185          190
181 Asn Glu
185 <210> SEQ ID NO: 3
187 <211> LENGTH: 218
189 <212> TYPE: DNA
191 <213> ORGANISM: Streptococcus agalactiae

```

RAW SEQUENCE LISTING

DATE: 10/30/2001

PATENT APPLICATION: US/09/868,352

TIME: 14:31:45

Input Set : A:\gje-70.txt

Output Set: N:\CRF3\10302001\I868352.raw

```

195 <220> FEATURE:
197 <221> NAME/KEY: CDS
199 <222> LOCATION: (1)..(216)
201 <223> OTHER INFORMATION:
205 <400> SEQUENCE: 3
206 atc aga gca tat tct ggt cct ctt tgc gtt ttc ctg cca cgt ttt aaa      48
207 ile arg ala tyr ser gly pro leu ser val phe leu pro arg phe lys
208 1          5          10          15
210 gct tgt gat ata ata gtc aat gtg agg agg act atc atg tta ttt aag      96
211 ala cys asp ile ile val asn val arg arg thr ile met leu phe lys
212          20          25          30
214 gaa aaa att cct gga cta ata tta tgc ttt att att gct ata cca tct      144
215 glu lys ile pro gly leu ile leu cys phe ile ile ala ile pro ser
216          35          40          45
218 tgg ttg ctt ggg ctt tat ctc cct tta ata gga gca cca gtc ttt gct      192
219 trp leu leu gly leu tyr leu pro leu ile gly ala pro val phe ala
220          50          55          60
222 atc ttg att gga ata att gtt gga tc      218
223 ile leu ile gly ile ile val gly
224 65          70
227 <210> SEQ ID NO: 4
229 <211> LENGTH: 72
231 <212> TYPE: PRT
233 <213> ORGANISM: Streptococcus agalactiae
237 <400> SEQUENCE: 4
239 ile arg ala tyr ser gly pro leu ser val phe leu pro arg phe lys
240 1          5          10          15
243 ala cys asp ile ile val asn val arg arg thr ile met leu phe lys
244          20          25          30
247 glu lys ile pro gly leu ile leu cys phe ile ile ala ile pro ser
248          35          40          45
251 trp leu leu gly leu tyr leu pro leu ile gly ala pro val phe ala
252          50          55          60
255 ile leu ile gly ile ile val gly
256 65          70
259 <210> SEQ ID NO: 5
261 <211> LENGTH: 705
263 <212> TYPE: DNA
265 <213> ORGANISM: Streptococcus agalactiae
269 <220> FEATURE:
271 <221> NAME/KEY: CDS
273 <222> LOCATION: (1)..(705)
275 <223> OTHER INFORMATION:
279 <400> SEQUENCE: 5
280 atg aat aaa aga aga aaa tta tca aaa ttg aat gta aaa aar caa cat      48
281 met asn lys arg arg lys leu ser lys leu asn val lys lys gln his
282 1          5          10          15
284 tta gct tat gga gct atc act tta gta gcc ctt ttt tca tgt att ttg      96
285 leu ala tyr gly ala ile thr leu val ala leu phe ser cys ile leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,352

DATE: 10/30/2001

TIME: 14:31:45

Input Set : A:\gje-70.txt

Output Set: N:\CRF3\10302001\I868352.raw

```

286          20          25          30
288 gct gta acg gtc atc ttt aaa agt tca caa gtt act act gaa tct ttg      144
289 Ala Val Thr Val Ile Phe Lys Ser Ser Gln Val Thr Thr Glu Ser Leu
290          35          40          45
292 tca aaa gca gat aaa gtt cgc gta gcc aaa aaa tca aaa atg act aag      192
293 Ser Lys Ala Asp Lys Val Arg Val Ala Lys Lys Ser Lys Met Thr Lys
294          50          55          60
296 gcg aca tct aaa tca aaa gta gaa gat gta aaa cag gct cca aaa cct      240
297 Ala Thr Ser Lys Ser Lys Val Glu Asp Val Lys Gln Ala Pro Lys Pro
298 65          70          75          80
300 tct cag gca tct aat gaa gcc cca aaa tca agt tct caa tct aca gaa      288
301 Ser Gln Ala Ser Asn Glu Ala Pro Lys Ser Ser Ser Gln Ser Thr Glu
302          85          90          95
304 gct aat tct cag caa caa gtt act gcg agt gaa gag acg gct gta gaa      336
305 Ala Asn Ser Gln Gln Gln Val Thr Ala Ser Glu Glu Thr Ala Val Glu
306          100          105          110
308 caa gca gtt gta aca gaa ata ccc ctg cta cca gtc agg cac aac aac      384
310 Gln Ala Val Val Thr Glu Ile Pro Leu Leu Pro Val Arg His Asn Asn
311          115          120          125
313 ctt tat gct gtt act gag aca cct tac aac cct gct caa cca cca gac      432
314 Leu Tyr Ala Val Thr Glu Thr Pro Tyr Asn Pro Ala Gln Pro Pro Asp
315          130          135          140
317 caa gtg gcc agg tat gag caa tgg aaa tac tgc cag gcg gtc gga tct      480
318 Gln Val Ala Arg Tyr Glu Gln Trp Lys Tyr Cys Gln Ala Val Gly Ser
319 145          150          155          160
321 gct gct gca gca caa atg gct gct gca aca gga gtc cct cag tct act      528
322 Ala Ala Ala Ala Gln Met Ala Ala Ala Thr Gly Val Pro Gln Ser Thr
323          165          170          175
325 tgg gaa cat att att gcc cgt gaa tca aat ggt aat cct aat gtt gct      576
326 Trp Glu His Ile Ala Arg Glu Ser Asn Gly Asn Pro Asn Val Ala
327          180          185          190
329 aat gcc tca gga gct tca gga ctt ttc caa acg atg cca ggt tgg ggt      624
330 Asn Ala Ser Gly Ala Ser Gly Leu Phe Gln Thr Met Pro Gly Trp Gly
331          195          200          205
333 tca aca gct aca gtt cag gat caa gta att cag cta tta aag ctt att      672
334 Ser Thr Ala Thr Val Gln Asp Gln Val Ile Gln Leu Leu Lys Leu Ile
335          210          215          220
337 cgt gct caa ggg tta tca gct ggg tac cag tga      705
338 Arg Ala Gln Gly Leu Ser Ala Gly Tyr Gln
339 225          230
342 <210> SEQ ID NO: 6
344 <211> LENGTH: 234
346 <212> TYPE: PRT
348 <213> ORGANISM: Streptococcus agalactiae
352 <400> SEQUENCE: 6
354 Met Asn Lys Arg Arg Lys Leu Ser Lys Leu Asn Val Lys Lys Gln His
355 1          5          10          15
358 Leu Ala Tyr Gly Ala Ile Thr Leu Val Ala Leu Phe Ser Cys Ile Leu
359          20          25          30

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/868,352

DATE: 10/30/2001

TIME: 14:31:45


Input Set : A:\gje-70.txt

Output Set: N:\CRF3\10302001\I868352.raw

```

362 Ala Val Thr Val Ile Phe Lys Ser Ser Gln Val Thr Thr Glu Ser Leu
363      35      40      45
366 Ser Lys Ala Asp Lys Val Arg Val Ala Lys Lys Ser Lys Met Thr Lys
367      50      55      60
370 Ala Thr Ser Lys Ser Lys Val Glu Asp Val Lys Gln Ala Pro Lys Pro
371 65      70      75      80
374 Ser Gln Ala Ser Asn Glu Ala Pro Lys Ser Ser Ser Gln Ser Thr Glu
375      85      90      95
378 Ala Asn Ser Gln Gln Gln Val Thr Ala Ser Glu Glu Thr Ala Val Glu
379      100     105     110
382 Gln Ala Val Val Thr Glu Ile Pro Leu Leu Pro Val Arg His Asn Asn
383      115     120     125
386 Leu Tyr Ala Val Thr Glu Thr Pro Tyr Asn Pro Ala Gln Pro Pro Asp
387      130     135     140
390 Gln Val Ala Arg Tyr Glu Gln Trp Lys Tyr Cys Gln Ala Val Gly Ser
391 145     150     155     160
394 Ala Ala Ala Ala Gln Met Ala Ala Ala Thr Gly Val Pro Gln Ser Thr
395      165     170     175
398 Trp Glu His Ile Ile Ala Arg Glu Ser Asn Gly Asn Pro Asn Val Ala
399      180     185     190
402 Asn Ala Ser Gly Ala Ser Gly Leu Phe Gln Thr Met Pro Gly Trp Gly
403      195     200     205
406 Ser Thr Ala Thr Val Gln Asp Gln Val Ile Gln Leu Leu Lys Leu Ile
407      210     215     220
410 Arg Ala Gln Gly Leu Ser Ala Gly Tyr Gln
411 225     230
414 <210> SEQ ID NO: 7
416 <211> LENGTH: 594
418 <212> TYPE: DNA
420 <213> ORGANISM: Streptococcus agalactiae
424 <220> FEATURE:
426 <221> NAME/KEY: CDS
428 <222> LOCATION: (1)..(594)
430 <223> OTHER INFORMATION:
434 <400> SEQUENCE: 7
435 atg act gaa cca ttt ttt gat aaa gaa tta act tgt cgc cca att gaa      48
436 Met Thr Glu Pro Phe Phe Asp Lys Glu Leu Thr Cys Arg Pro Ile Glu
437 1      5      10      15
439 gcc att cct gaa ttg ttg gaa ttc gat att acc gtt cgt gga gac aac      96
440 Ala Ile Pro Glu Leu Leu Glu Phe Asp Ile Thr Val Arg Gly Asp Asn
441      20      25      30
443 cgt gga tgg ttc aaa gag aac ttt caa aaa gaa aaa atg ata ccg ctt      144
444 Arg Gly Trp Phe Lys Glu Asn Phe Gln Lys Glu Lys Met Ile Pro Leu
445      35      40      45
447 ggt ttc cca gaa agc ttc ttt gag gca gac aaa cta caa aat aat att      192
448 Gly Phe Pro Glu Ser Phe Phe Glu Ala Asp Lys Leu Gln Asn Asn Ile
449      50      55      60
451 tcg ttt aca aaa aaa aat act ttg cga ggt ctc cat gca gag cct tgg      240
452 Ser Phe Thr Lys Lys Asn Thr Leu Arg Gly Leu His Ala Glu Pro Trp

```


 Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/868,352

DATE: 10/30/2001

TIME: 14:31:46

Input Set : A:\gje-70.txt

Output Set: N:\CRF3\10302001\I868352.raw

L:643 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9
L:2178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:2302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.